This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/705, C07H 21/04, C12N 1/21, 5/16, 15/12, 15/63, C12P 21/00, C12Q 1/25, 1/68, G01N 33/53

A1

(11) International Publication Number: WO 96/21677

(43) International Publication Date: 18 July 1996 (18.07.96)

(21) International Application Number: PCT/US95/16311

(22) International Filing Date: 8 December 1995 (08.12.95)

(30) Priority Data:
08/372,652 13 January 1995 (13.01.95) US

(71) Applicant: THE GENERAL HOSPITAL CORPORATION [US/US]; 55 Fruit Street, Boston, MA 02114 (US).

(72) Inventors: MOORE, David; One Isaac Sprague Drive, Hingham, MA 02043 (US). SEOL, Wongi; 812 Memorial Drive, A1105, Cambridge, MA 02139 (US). CHOI, Hueng-Sik; Korea Research Institute of Bioscience & Biotechnology, Immunology Division, Molecular Biomedicine Group, Taejon (KR).

(74) Agent: LECH, Karen, F.; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).

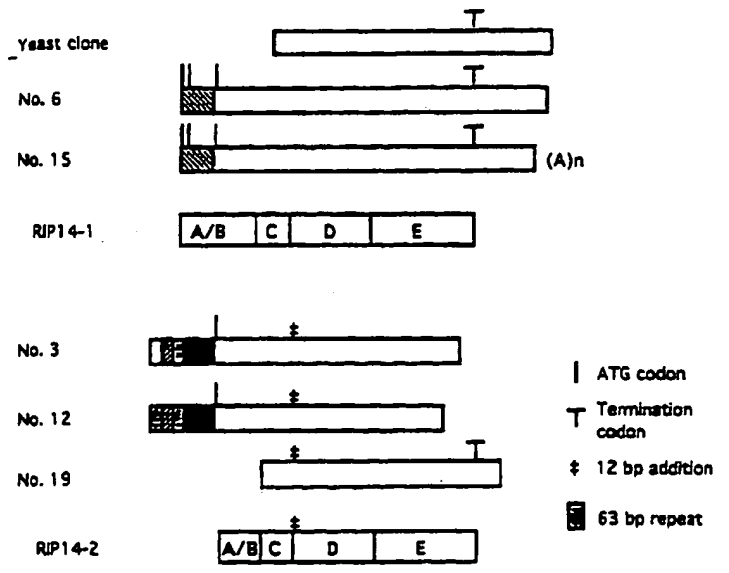
(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS

(57) Abstract

Disclosed is a method for determining whether a test protein is capable of interacting with a retinoid X receptor protein. The method involves: (a) providing a host cell which contains (i) a reporter gene operably linked to a protein binding site; (ii) a first fusion gene which expresses a first fusion protein, the first fusion protein including a retinoid X receptor protein covalently bonded to a binding moiety which is capable of specifically binding to the protein binding site; and (iii) a second fusion gene which expresses a second fusion protein, the second fusion protein including the test protein covalently bonded to a gene activating moiety; and (b) determining whether the test protein increases expression of the reporter gene as an indication of its ability to interact with the retinoid X receptor protein. Also disclosed is purified DNA encoding retinoid X receptor-interacting proteins and the polypeptides expressed from such DNA.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

RETINOID X RECEPTOR-INTERACTING POLYPEPTIDES AND RELATED MOLECULES AND METHODS

Background of the Invention

This invention relates to receptor proteins.

This invention was made in part with Government funding, and the Government therefore has certain rights in the invention.

The retinoid X receptors (RXRs) are members of a large superfamily of intracellular hormone receptors. These proteins bind to specific DNA sequences and directly regulate transcription of target genes in response to activation by their specific ligands (Leid et al., Trends Biochem. Sci. 17:427-433, 1992; Leid et al., Cell 68:377-395, 1992; Mangelsdorf et al., Nature 345:224-229, 1990 and Yu et al., Cell 67:1251-1266, 1991). The RXRs belong to a large subgroup of the superfamily defined by a conserved subregion within the DNA binding domain. This group also includes the receptors for retinoic acid, thyroid hormone, and vitamin D as well as a number of other less well characterized proteins, called orphan receptors, that do not have known ligands. As monomers, the members of this class can bind to sequences related to the hexameric consensus AGGTCA. RXR homodimers bind to tandem repeats of this consensus separated by a single base pair (Manglesdorf et al., Cell 66:555-561, 1991), and apparently to additional elements including β-RARE (Zhang et al., Nature 358:587-591, 1992). These homodimer binding sites confer specific response to 9-*cis*-RA (9-*cis*-RA), the ligand for the RXRs. In addition, the RXRs heterodimerize with a variety of other family members, including the receptors for all-*trans*-retinoic acid, thyroid hormone (T3), and vitamin D. This heterodimerization strongly increases the affinity of these receptors for their specific response elements

(Yu et al., Cell 67:1251-1266, 1991; Zhang et al., Nature 358:587-591, 1992; Bugge et al., EMBO J. 11:1409-1418, 1992), and recent evidence also demonstrates that it is also required for full hormone dependent transcriptional activity of at least the thyroid hormone receptor-RXR complex.

Mammals have three genes encoding α, β, and γ isoforms of RXR (Mangelsdorf et al., Genes & Dev. 6:329-344, 1992). The expression patterns of murine RXRs (Mangelsdorf et al., Genes & Dev. 6:329-344, 1992) and homologues of RXR found in *Xenopus* (Blumberg et al., Proc. Natl. Acad. Sci. USA 89:2321-2325, 1992) and *Drosophila* (Oro et al., Nature 347:298-301, 1990) suggest that the members of the RXR family play important roles in several aspects of development and central nervous system differentiation as well as in adult physiology. Based on both their specific response to the 9-*cis*-RA metabolite and their heterodimerization with the RARs, it is clear that the RXRs play a central role in the broad regulatory effects of retinoids. Moreover, their heterodimeric interactions with other family members indicate that the RXRs also play a central role in response to thyroid hormone, vitamin D, and perhaps other compounds. This dual function is unique within the nuclear receptor superfamily.

Summary of the Invention

In a first aspect, the invention generally features a method for determining whether a test protein is capable of interacting with a retinoid X receptor (RXR) protein. The method involves: (a) providing a host cell which contains (i) a reporter gene operably linked to a protein binding site; (ii) a first fusion gene which expresses a first fusion protein, the first fusion protein including a retinoid X receptor protein

covalently bonded to a binding moiety which is capable of specifically binding to the protein binding site; and (iii) a second fusion gene which expresses a second fusion protein, the second fusion protein including the test protein covalently bonded to a gene activating moiety; and (b) determining whether the test protein increases expression of the reporter gene as an indication of its ability to interact with the retinoid X receptor protein.

In a preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor (preferably, 9-*cis*-RA) and identifying a ligand-dependent interacting protein by its ability to increase expression of the reporter gene upon treatment of the cell by the ligand. In another preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor and identifying a ligand-independent interacting protein by its ability to increase expression of the reporter gene both in the presence and in the absence of ligand treatment. In yet another preferred embodiment, the method further involves treating the host cell with a ligand which binds the retinoid X receptor and identifying a ligand-sensitive interacting protein by its ability to increase expression of the reporter gene in the absence but not in the presence of ligand treatment.

In other preferred embodiments, the gene activating moiety is the gene activating moiety of B42.

In a second aspect, the invention features a substantially pure preparation of a retinoid X receptor (RXR)-interacting protein. Preferably, the RXR-interacting protein is RIP14, RIP15, RIP110, or RIP13; or includes an amino acid sequence substantially identical to an amino acid sequence shown in any of Figs. 4, 5, 10,

and 11 (SEQ ID NOS: 1-5); is derived from a mammal, for example, a human; binds a β-RARE site in the presence of RXR; or binds an EcRE site in the presence of RXR.

The invention also features purified DNA (for example, cDNA) which includes a sequence encoding an RXR-interacting protein, preferably encoding a human RXR-interacting protein (for example, the RXR-interacting proteins RIP14 (SEQ ID NO: 6, 14), RIP15 (SEQ ID NO: 7), RIP110 (SEQ ID NO: 8), or RIP13 (SEQ ID NO: 9)); a vector and a cell which includes a purified DNA of the invention; and a method of producing a recombinant RXR-interacting protein involving providing a cell transformed with DNA encoding an RXR-interacting protein positioned for expression in the cell; culturing the transformed cell under conditions for expressing the DNA; and isolating the recombinant RXR-interacting protein. The invention further features recombinant RXR-interacting protein produced by such expression of a purified DNA of the invention.

As used herein, "reporter gene" is meant a gene whose expression may be assayed; such genes include, without limitation, *lacZ*, amino acid biosynthetic genes, e.g. the yeast *LEU2* gene, luciferase, or the mammalian chloramphenicol transacetylase (CAT) gene. Reporter genes may be integrated into the chromosome or may be carried on autonomously replicating plasmids (e.g., yeast 2μ plasmids).

By "operably linked" is meant that a gene and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins or proteins which include transcriptional activation domains) are bound to the regulatory sequence(s).

By a "binding moiety" is meant a stretch of amino acids which is capable of directing specific polypeptide

binding to a particular DNA sequence (i.e., a "protein binding site"). LexA represents a preferred DNA binding moiety in the invention. However, any other transcriptionally-inert or essentially transcriptionally-inert DNA binding domain may be substituted. The GAL4 DNA binding domain represents a somewhat less preferred DNA binding moiety for the system described herein.

By "gene activating moiety" is meant a stretch of amino acids which is capable of inducing the expression of a gene to whose control region it is bound. As used herein, a "weak gene activating moiety" is meant a stretch of amino acids which induces gene expression at a level below the level of activation effected by GAL4 activation region II (Ma and Ptashne, Cell 48:847, 1987) and is preferably at or below the level of activation effected by the B42 activation domain of Ma and Ptashne (Cell 51:113, 1987). Levels of activation may be measured using any downstream reporter gene system and comparing, in parallel assays, the level of expression stimulated by the GAL4- or B42-polypeptide with the level of expression stimulated by the polypeptide to be tested.

By "RXR-interacting protein" is meant a polypeptide which directly or indirectly physically interacts with a retinoid X receptor in the *in vivo* protein interaction assay described herein. Such an interaction may be hormone (or ligand) dependent or independent or may be hormone (or ligand) sensitive; it may also be transient in nature so long as the interaction is capable of producing a positive result in the interaction assay described herein. Preferably, such a polypeptide has an amino acid sequence which is at least 85%, preferably 90%, and most preferably 95% or even 99% identical to the amino acid sequence of an interacting protein described herein (e.g., RIP14, RIP15, RIP110, or RIP13) at the point of interaction with the

retinoid X receptor, or at least 80% and preferably 90% identical overall.

By "substantially pure" is meant a preparation which is at least 60% by weight (dry weight) the compound of interest, i.e., an RXR-interacting protein. Preferably the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate method, e.g., column chromatography, polyacrylamide gel electrophoresis, or HPLC analysis.

By "purified DNA" is meant DNA that is not immediately contiguous with both of the coding sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence.

By "substantially identical" is meant an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the function of the protein (assayed, e.g., as described herein). Preferably, such a sequence is at least 85%, more preferably 90%, and most preferably 95%

identical at the amino acid level to one of the sequences of Figs. 4, 5, 10, and 11 (SEQ ID NOS: 1-5). A "substantially identical" nucleic acid sequence codes for a substantially identical amino acid sequence as defined above.

By "transformed cell" is meant a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding (as used herein) an RXR-interacting protein.

By "positioned for expression" is meant that the DNA molecule is positioned adjacent to a DNA sequence which directs transcription and translation of the sequence (i.e., facilitates the production of, e.g., an RXR-interacting protein).

By "purified antibody" is meant antibody which is at least 60%, by weight, free from the proteins and naturally-occurring organic molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, antibody, e.g., RXR-interacting protein-specific antibody. A purified RXR-interacting protein antibody may be obtained, for example, by affinity chromatography using recombinantly-produced RXR-interacting protein and standard techniques.

By "specifically binds" is meant an antibody which recognizes and binds RXR-interacting protein but which does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which naturally includes RXR-interacting protein.

Other features and advantages of the invention will be apparent from the following detailed description thereof, and from the claims.

Description of the Drawings

The drawings are first described.

Figure 1 is a graph showing β-galactosidase expression conferred to LexA-RXR by B42-RIP clones. A yeast strain containing a *lacZ* reporter gene under the control of LexA binding sites and LexA-RXR was transformed with the indicated B42-fusion protein expression vectors. Strains coexpressing LexA-RXR and the indicated B42 fusions were grown overnight in liquid in the presence or absence of 10^{-6}M 9-*cis*-RA, added at the time of inoculation, and β-galactosidase activity was assayed.

Figure 2 is a Northern analysis of RIP14 (panel A) and RIP15 (panel B). A Northern blot containing mRNAs from the indicated tissues (Clontech, Palo Alto, CA) was hybridized with RIP14 and RIP15 probes. In a longer exposure of RIP14, several larger bands not evident in this exposure were observed in liver and kidney.

Figure 3 is a schematic representation of the structure of the RIP14 cDNA and protein isoforms. The original yeast clone isolated from the interaction trap is diagrammed at the top. The positions of the putative initiation and termination codons are indicated, and the 12 base pair (bp) addition in the D domain of RIP14-2 is shown as ‡. Various sequences specific to each isoform and a 63 bp repeat present in Clone 12, but not Clone 3 are indicated.

Figure 4 is the nucleotide (SEQ ID NO: 6, 14) and deduced amino acid sequences (SEQ ID NOS: 1, 2) of the RIP14 clone. Numbers of nucleotides and amino acids of each sequence are shown on the left side. For RIP14-1 (SEQ ID NO: 1), the sequence of Clone 6 is shown. Four amino acids unique to RIP14-2 (SEQ ID NO: 2) are also shown. The DNA binding domain (C domain) and a poly A signal are underlined. N-terminus unique to isoform 1 is

shown as italicized letters and the putative initiation codons as bold letters. Because the 5' ends of the RIP14-2 clones are different, both Clone 3 and Clone 12 are shown (SEQ ID NOS: 15-17). A 63 base pair direct repeat which occurs in the 5' end of Clone 12 is indicated as underlined and in italicized letters. The GenBank submission numbers of RIP14-1 is U09416. The submission number of RIP14-2 Clones 3 and 12 are U09417 and U09418, respectively.

Figure 5 is the nucleotide (SEQ ID NO: 7) and deduced amino acid sequences (SEQ ID NO: 3) of RIP15. The in frame termination codon in front of the initiation codon, the DNA binding domain (C domain), and a poly A signal are underlined. The GenBank submission number of this sequence is U09419.

Figure 6 is a photograph showing an SDS-polyacrylamide gel analysis of *in vitro* translated RIP proteins. Lane 1: RIP14-1, 2: RIP14-2 (No. 3), 3: RIP14-2 (No. 12), 4: RIP15, 5: positive control (luciferase, MW *ca* 69 Kd). Clone No. 15 was used for expression of RIP14-1, and the full length derivatives of both Clones No. 3 and 12 (as described herein) for RIP14-2. Molecular weight markers are shown.

Figure 7 is a schematic diagram showing a comparison of RIP14 and RIP15 sequences to other receptors. Percent sequence identities of RIP14 (panel A) and RIP15 (panel B) compared to DNA and ligand binding domains of the indicated receptor superfamily members are shown. For comparison, the isoform of each receptor member showing the highest score in a GenBank database search was used. TR: mouse thyroid hormone receptor α-1 (accession number: P16416); RAR: zebrafish retinoic acid receptor γ (accession number: L03400); RXR: mouse retinoic acid X receptor α (accession number: P28700); EcR: *Drosophila* ecdysone receptor (accession number:

P34021). Lengths of the various domains of the receptors are not to scale.

Figure 8 is a series of photographs showing DNA binding by the RXR-interacting proteins. In panel A, RIP14-1 without (lane 2) or with RXR (lanes 3-5) was incubated with the hsp27 promoter (EcRE) end-labeled with [^{32}P]ATP. In panel B, RIP15 (lanes 2-5) and RIP14-1 (lanes 9-12) without (lanes 2 and 9) or with (lanes 3-5 and 10-12) RXR were incubated with βRARE end-labeled with [^{32}P]ATP. In both cases, a 50-fold molar excess of unlabeled specific (sp; EcRE, lane A4 and βRARE, lanes B4 and B11) or nonspecific (ns; AP1, lanes A5, B5, and B12) oligomers were added with the labeled probe. Incubations with probe alone are shown in lanes A1, B1, and B8. Cell lysates used for *in vitro* translation were also incubated with the probes without (lanes A6 and B6) or with (lanes A7 and B7) RXR. RIP14-2 and the RIP14-1 chimeras containing a four amino acid insertion in the D domain (RIP14C) were also incubated with βRARE and RXR as shown in lanes B13 and B14, respectively. Specific bands are indicated by arrows.

Figure 9 is a graph showing activity of RIP14-1, RIP14-2, and RIP-15 in mammalian cells. Vectors expressing intact RIPs, RXR, or the CDM vector were cotransfected into HepG2 cells as indicated herein with a luciferase reporter plasmid containing three copies of the β-RARE and pTKGH as an internal control. 9-*cis*-RA was added at 10^{-6}M approximately 18 hours after transfection. Results are luciferase expression normalized to the hGH internal control. Consistent results were obtained in three independent experiments.

Figure 10 is the nucleotide (SEQ ID NO: 8) and deduced amino acid sequences (SEQ ID NO: 4) of RIP110.

Figure 11 is the nucleotide (SEQ ID NO: 9) and deduced amino acid sequences (SEQ ID NO: 5) of RIP13.

Detailed Description

Applicants have used an in vivo interaction trap system to identify and isolate proteins that physically interact with retinoid X receptors and, in particular, with the ligand binding domain of the human RXRα. These proteins are termed RXR-Interacting Proteins (or RIPs). The isolation and characterization of exemplary RIPs now follows.

Isolation of proteins that interact specifically with RXR

Recently, several genetic methods have been used to identify and characterize protein-protein interactions (e.g., Fields et al., Nature 340:245-246, 1989; Gyuris et al., Cell 75:791-803, 1993). The principal idea of these systems is that transcription activation and DNA binding are quite distinct functions within most eukaryotic transcription activators, generally localized to two separate domains. Many functional examples of chimeric transcriptional activators consisting of the DNA binding domain of one protein attached to a heterologous activation domain have been characterized (Green et al., Nature 325:75-78, 1987; Ma et al., Cell 51:113-119, 1987). The fact that this attachment can be indirect, mediated by protein-protein interaction rather than the covalent linkage of separate domains of a single protein forms the basis of the selection. A version of this type of system, called the interaction trap, has allowed the isolation of several new proteins that interact with several different targets, including Max (Zervos et al., Cell 72:223-232, 1993), Cdc2 (Gyuris et al., Cell 75:791-803, 1993), and RAG-1 (Coumo et al., Proc. Natl. Acad. Sci. USA in press., 1994).

We used the interaction trap (Gyuris et al., Cell 75:791-803, 1993) to identify cDNAs encoding proteins that interact with the ligand binding domain of human RXRα. As shown in Figure 1, a chimeric protein

consisting of the hinge (D) and ligand binding (E) domains of RXRα fused to the intact bacterial LexA repressor protein (LexA-RXR) is not a strong transcriptional activator in yeast, in either the presence or absence of 9-*cis*-RA. However, LexA-RXR does activate expression from LexA binding sites in cells which also express a fusion protein consisting of a transcriptional activation domain joined to another protein, such as the thyroid hormone receptor, which interacts specifically with RXR.

To isolate RXR-interacting proteins, we constructed a mouse liver cDNA library in a derivative of the yeast vector pJG4-5 (Gyuris et al., Cell 75:791-803, 1993; see below), in which the cDNA sequences are fused to the B42 transcriptional activation domain (Ma et al., Cell 51:113-119, 1987). The liver was chosen because it is a major target organ for the actions of a number of nuclear hormone receptor superfamily members. This library was introduced, as described herein, into a LexA-RXR expressing host in which transcription of both the β-galactosidase (β-gal) and LEU2 genes was under the control of LexA binding sites.

From $3x10^6$ primary yeast transformants, a number of β-gal expressing colonies were identified in two independent screens carried out in the presence or absence of 9-*cis*-RA. Since expression of the B42-cDNA fusion proteins was induced by growth on galactose, candidates obtained from each condition were tested for galactose-dependent expression of both β-gal and LEU2, in the presence or absence of 9-*cis*-RA, using appropriate indicator plates (see below). To test the specificity of the interaction with RXR, cDNA plasmids were rescued from a number of candidates that showed appropriate galactose dependency and reintroduced into hosts expressing LexA alone or other LexA fusion proteins (e.g., LexA-Cdc2).

Candidates showing specific interaction with LexA-RXR were sequenced across the B42 fusion junction using an appropriate vector primer, and additional sequence was determined. The deduced amino acid sequences were compared to those in the GenBank and EMBL databases using the GCG (Devereux et al., Nucleic Acids Res. 12:387-395, 1984) and BLAST programs (Altschul et al., J. Mol. Biol. 215:403-410, 1990). This sequence comparison demonstrated that a number of the independently isolated clones encoded known RXR heterodimer partners, either PPAR (eight clones) or RARα (six clones), providing strong confirmation of the specificity of the screening. All of these clones included intact ligand binding domains, as expected from the fact that the major heterodimerization function lies within that domain. For PPAR (Issemann et al., Nature 347:645-650, 1990), three clones started at amino acid 84, just N-terminal to the DNA binding (C) domain, four others started at 91, within the C domain, and one at 170, just past the C domain. For RAR (Leroy et al., EMBO J. 10:59-69), all six clones started at amino acid 132, within the C domain. Although previous results demonstrated that LexA-RXR interacts strongly with a chimera consisting of B42 fused to the thyroid hormone receptor, no B42-TR clones were obtained, presumably as a consequence of the very low level of expression of the TR mRNAs in liver.

Six clones representing three independent B42 fusions to the vitamin D binding protein (Yang et al., Genomics 7:509-516, 1990) were isolated. Since this secretory protein is unlikely to interact with the nuclear RXR protein in intact cells, it is difficult to explain why these clones were obtained. It is possible that the vitamin D binding protein shares some structural similarity with the vitamin D receptor, which is an RXR heterodimer partner, and that this structural

conservation is the basis for the interaction. However, it is also possible that the interaction of this protein with RXR is simply an artifact of the yeast system, and these clones have not been studied further.

Several additional clones encoded novel proteins. Two, RIP14 and RIP15, were previously undescribed orphan members of the nuclear receptor superfamily. As with the PPAR and RAR isolates, the B42 fusion junctions in both cases were near the beginning of the hinge (D) domain that separates the DNA (C) and ligand binding (E) domains. Two other clones, RIP13 and RIP110, showed no significant similarity to any known protein and are candidate transcriptional co-activators.

The level of β-gal expression conferred by several of the RXR interactors was examined more quantitatively (as described herein). The results of β-galactosidase assays of liquid cultures of a series of appropriate strains are shown in Figure 1. As expected from previous results with B42-TR and numerous biochemical studies (Leid et al., Cell 68:377-395, 1992; Zhang et al., Nature 358:587-591, 1992), the interaction of B42-RAR with LexA-RXR was independent of the presence or absence of 9-*cis*-RA. Cells coexpressing LexA-RXR and the RIP14 and RIP15 chimeras showed levels of β-gal expression comparable to B42-RAR regardless of the presence or absence of the RXR ligand 9-*cis*-RA, indicating a relatively strong, ligand independent interaction with RXR. With RIP13, the significant β-gal expression observed in the absence of 9-*cis*-RA was increased approximately three fold in the presence of the ligand. For RIP110, only basal levels of expression were observed in the absence of 9-*cis*-RA, but the level of β-gal expression was strongly induced in the presence of 9-*cis*-RA, indicating that the interaction of this protein with RXR is dependent on the presence of ligand.

Human cDNAs encoding any of these RIP polypeptides may be isolated using human cDNA libraries (for example, a human liver cDNA library) and standard techniques of hybridization.

Interaction of RIPs with other Receptors

Interactions of the RIP clones with other superfamily members, including RAR, TR, the glucocorticoid receptor (GR), and MB67, an orphan isolated in this lab (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994) were also tested using a series of appropriate LexA fusions. In particular, the experiments shown in Table 1 were carried out as follows. Yeast transformants containing a *lacZ* reporter gene under the control of LexA binding sites and the indicated B42- and Lex-fusion protein expression vectors were transferred onto galactose-Ura$^-$His$^-$Trp$^-$ plates containing X-gal and incubated for two days. Relative levels of β-galactosidase activity were estimated and are depicted as follows: B, blue (strong interaction); LB, light blue (weak interaction); W, white (no interaction); nt, not tested. At least three separate colonies were tested on the same plate for interaction. To test the effects of ligands, 100μl of a 10^{-6}M solution of the appropriate ligand was spread onto plates just before inoculation of cells (9-*cis*-RA for RXR, T3 for TR, and all-*trans*-RA for RAR). For B42-PPAR, the plasmid isolated from the screening was used. Both full length and truncated RAR fusions to LexA were tested with identical results for all the B42 fusions except B42-110, which was tested only with the full length fusion.

As indicated in Table 1, RIP13 and RIP110 interacted with all proteins except glucocorticoid receptor (GR), whereas RIP14 and RIP15 interacted with RXR only.

Table 1

Lex Fusion	RXR		TR		RAR		MB67		GR
ligand	+	-	+	-	+	-			
B42 fusion									
13	B	B	B	B	B	B	B		W
14	B	B	W	W	W	W	W		W
15	B	B	W	W	W	W	W		W
110	B	LB	B	W	B	B	B		W
PPAR	B	B	W	W	nt	nt	nt		nt
RXR	B	B	B	B	B	B	B		W

RIP13's interaction with all of the conventional receptors remained unaffected by the presence or absence of ligands. As with LexA-RXR, RIP110 interacted with LexA-TR in a ligand-dependent manner. Its interaction with LexA-RAR was not dependent on retinoic acid, however, and it also interacted constitutively with MB67, which does not have known ligands. This orphan showed substantial transcriptional activity in mammalian cells grown in the absence of any exogenously added ligands. The unique character of RIP13 and RIP110 in these interactions suggests that these proteins may have very important roles in conserved functions of the nuclear hormone receptor superfamily. In particular, the interaction of RIP13 with a wide array of superfamily members indicates that it may be involved in any of several processes that are common to a number of different receptors. In addition to direct transcriptional regulation, this could include nuclear transport, for example. The ligand-dependent nature of the interaction of RIP110 suggests that it may be directly involved in ligand-dependent transcription or other activities.

Isolation of Full-Length RIP14 and RIP15 cDNAs

Northern blot analysis as described herein revealed that RIP14 is expressed only in liver and kidney in mouse, predominantly as a broad band of approximately 1.8 to 2.2 kb (Figure 2). Much lesser amounts of three or four higher molecular weight species were also observed. In contrast, an approximately 2.3 kb RIP15 mRNA was universally expressed in a number of tissues. To obtain full-length clones of these mRNAs, a mouse liver cDNA library was constructed and screened by conventional hybridization with RIP14 and RIP15 probes. Eight separate clones were obtained for RIP14, and four for RIP15.

All eight RIP14 clones were analyzed by digestion with multiple restriction enzymes and either partial or complete sequencing. As diagrammed in Figure 3, these clones may be divided into two subgroups that apparently encode distinct isoforms, referred to as RIP14-1 and RIP14-2. Based on initiation at the first methionine of the open reading frame, the RIP14-1 isoform is a protein of 484 amino acids. Since there are no in frame termination codons upstream from this methionine, it remains possible that this isoform includes additional N-terminal sequence. The RIP14-1 mRNA, however, is approximately 1.8 to 2.0 kb, as demonstrated by the hybridization of an oligonucleotide probe specific for the 5′ end of the RIP14-1 sequence to only the lower portion of the broad band recognized by a common probe, and the cloned sequences account for approximately 2 kb, assuming a 200 nucleotide poly A tail. Thus, if such an N-terminal extension exists, it must be minimal. As many members of the receptor superfamily have 5′ untranslated regions of several hundred nucleotides that include additional upstream AUG codons, it is also possible that the correct RIP14-1 sequence begins with a methionine

downstream of that indicated in Figure 4. For simplicity, we assume that the indicated reading frame is full length. As described below, results from *in vitro* translation are consistent with this assignment.

The RIP14-2 group differed from RIP14-1 in two respects. First, Clones 3 and 12 shared related 5′ sequences that diverged from those of RIP14-1 at a position within the apparent RIP14-1 coding region. Most of the 5′ sequence specific to clone 12 consisted of an additional copy of a 63 base pair sequence from the region common to the RIP14-2 class (Figures 3 and 4). The basis for this variation is not known. As expected, hybridization with a probe specific for these more extensive RIP14-2 sequences detected the upper portion of the broad band recognized by the common probe, indicating that the full length RIP14-2 mRNA is approximately 2.0 to 2.2. kb. This corresponds well to the approximately 2.2 kb predicted from the sequence obtained, plus a poly A tail. Because of the divergent sequence, the first methionine of the open reading frame of the 451 amino acid RIP14-2 isoform corresponds to residue 38 of the predicted RIP14-1 sequence. RIP14-2 also differed from RIP14-1 by an insertion of four amino acids located four amino acids downstream of the C-terminus of the DNA binding domain. Although this insertion is only observed in clones that include the RIP14-2 type 5′ end, it is possible that it is also present in a subset of mRNAs of the RIP14-1 type, and that it may be absent in a subset of mRNAs of the RIP14-2 type. However, there is currently no evidence for the existence of these two potential additional isoforms.

In contrast to the complex structure of RIP14, all of the clones of RIP15 show identical sequence, except one that is 64 base pairs longer at the 5′ end. All four clones contain the same open reading frame for the

predicted 446 amino acid RIP15 protein, with an in-frame terminator nine bases upstream of the initiation codon (Figure 5).

In vitro translation of transcripts for the two RIP14 isoforms produced a slightly bigger protein for RIP14-1 than for RIP14-2, approximately 57 and 55 Kd (Figure 6). This corresponds well to the predicted sizes of approximately 55 and 52 Kd, and is consistent with the assignment of the initiator codons. The calculated molecular weight of RIP15 is approximately 48 Kd. However, the major *in vitro* translation product was approximately 60 Kd (Figure 6). The basis for this apparently aberrant migration is unclear.

In sum, therefore, the RIP15 gene apparently encodes a single product, while the RIP14 gene expresses at least two closely related isoforms. By analogy with several examples of alternate promoter utilization for other superfamily members, it seems likely that this mechanism accounts for the distinct 5′ ends of the two major classes of RIP14 cDNAs. Because the cDNAs appear to be full length, the alternative possibility that they are derived by alternative splicing of a common transcript from a single promoter seems less likely. In contrast to this relatively common N-terminal variation, the variation of the RIP14 isoforms within the D domain is apparently unique within the receptor superfamily. Without information on the structure of the RIP14 gene, it is not possible to be certain how this variation is generated. However, since the first 6 nucleotides of the inserted sequence match the consensus 5′ or donor splice site (GU(A/G)AGU) (SEQ ID NO: 10), it is quite likely that it is a consequence of alternative donor site utilization.

In Figure 7, the amino acid sequences of the two orphans were compared to those of several other members

of the nuclear hormone receptor superfamily. In the DNA binding (C) domain, RIP14 was most closely related to insect ecdysone receptors, sharing 82% sequence identity with that of *Drosophila* (Koelle et al., Cell 67:59-77, 1991), for example. Interestingly, RIP15 is the superfamily member next most closely related to RIP14 in this domain, with 67% identity. The DNA binding domains of RIP15 and the ecdysone receptors share 64% identity, and these three sequences form a rather divergent subgroup within the superfamily. A distinctive feature of the RIP15 C domain sequence is an insertion of two amino acids in the short region between the two zinc modules. Although they do not show particularly strong overall sequence identity with the RIP14/RIP15/ecdysone receptor subgroup elsewhere in the C domain, a similar insertion is present in the thyroid hormone receptors.

Both RIP14 and RIP15 included matches to all of the conserved sequence motifs present in the putative ligand binding and dimerization (E) domains of other orphans and conventional receptors (Seagraves et al., Genes & Dev. 4:204-219, 1990; Amero et al., Mol. Endocrinol 6:3-8, 1992; Laudet et al., EMBO J 11:1003-1013, 1992), including a conserved C-terminal sequence associated with ligand dependent transcriptional activation (Danielian et al., EMBO J. 11:1025-1033, 1992). As with the C domain, overall comparisons based on ligand/dimerization domains placed both RIP14 and RIP15 in a divergent subgroup that also includes the ecdysone receptors. Within this region, RIP14 shares 42% and 37% identity with RIP15 and the ecdysone receptor, respectively, while RIP15 shares 42% identity with the ecdysone receptor. Overall, these three proteins are approximately as closely related to each other as the TRs are to the RARs.

Heterodimers of RIP14 and RIP15 with RXR Bind DNA Specifically

To identify DNA sequences recognized by each orphan, *in vitro* translated proteins were used for gel shift assays. Because of their strong sequence similarity to the ecdysone receptor, a well studied ecdysone response element (EcRE) from the *Drosophila* hsp27 promoter (Riddilough et al., EMBO J. 6:3729-3734, 1987) was tested for binding to both orphans, in the presence and absence of RXR. This element consists of two hexamers that match the receptor binding consensus AGGTCA, arranged as an inverted repeat separated by one base pair (IR-1). As demonstrated in Figure 8, panel A, RIP14-1 bound to the EcRE, but only in the presence of RXR. The binding of the RIP14-2 isoform to this element was weaker than that of RIP14-1 when similar amounts of RIP14 proteins were used. RIP15 did not bind to the EcRE regardless of the presence or absence of RXR. Several other DNA elements were tested in the gel shift assay, including the retinoic acid response element from the promoter of the human RARβ2 isoform (βRARE) (de The et al., Nature 343:177-180, 1990). In the presence of RXR, the βRARE was bound by both the RIP14 isoforms and RIP15 (Figure 8, panel B). Again, the binding of the RIP14-2/RXR heterodimer was weaker than the RIP14-1/RXR heterodimer when a similar amount of RIP14 proteins were used (Figure 8, panel B, lanes 10 and 14). In contrast to the results with the EcRE, RIP14-1 showed some binding to the βRARE in the absence of RXR.

The difference in apparent binding affinity between RIP14-1 and RIP14-2 could be due to either the variation in the short A/B domains or the four additional amino acids in the D domain of RIP14-2. The former would be consistent with the recently reported effects of different A/B domains on DNA binding by isoforms of the

orphan ROR (Giguere et al., Genes & Dev. 8:538-553, 1994). The latter would be consistent with the fact that the insertion occurs within a region called the T box (Wilson et al., Science 2546:107-110, 1992), which is associated with effects on DNA binding to direct repeats by heterodimers of other superfamily members. A chimeric receptor containing the four additional amino acids (MYTG) of RIP14-2 at the corresponding region of the D domain of RIP14-1 was constructed and tested for binding to both the βRARE and EcRE in the presence of RXR. As shown in Figure 8, panel B, lane 13, the binding of this chimeric protein (RIP14C) to the βRARE was similar to that of isoform 1 rather than isoform 2. The EcRE also showed a similar result. The lack of an effect of the insertion on RIP14-1 binding suggests that the differences in the A/B domain may determine relative binding affinity of the two isoforms.

From the above results, we conclude that both RIP14 and RIP15 bind to an overlapping set of specific elements as heterodimers with RXR.

The at least partially overlapping DNA binding specificity of RIP14 and RIP15 is consistent with the similarity of their DNA binding domains and suggests that they may have overlapping functional roles. Since both interact with the βRARE, it is possible that these functions include effects on the complex response to retinoids. However, the inactivity of both intact and chimeric versions of the two orphans in transient transfections indicates that both require activation by binding of as yet unidentified ligands or by other processes.

Function of RIP14 and RIP15 *in vivo*

To test the transcriptional activity of RIP14-1, RIP14-2, and RIP15, vectors expressing each were cotransfected into HepG2 cells with a luciferase reporter

plasmid in which three copies of the βRARE were inserted upstream of the TK promoter (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994; described herein). Expression from this reporter was transactivated more than 100 fold by RAR in the presence of retinoic acid, and 20-50 fold by the apparently constitutive orphan MB67 (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994). Neither of the two RIP14 isoforms or RIP15 was able to transactivate the βRARE reporter under any condition tested (Figure 9). This was confirmed using chimeras in which the A/B and DNA binding (C) domains of the thyroid hormone receptor (TR) were fused to the hinge (D) and ligand binding (E) domain of each orphan. In cotransfections of these chimeras with an analogous reporter plasmid containing two copies of the synthetic palindromic T3 response element (TREpal) (Brent et al., Mol. Endocrinol. 3:1996-2004, 1989), the TR-RIP14 chimera was not significantly different from the CDM vector alone. A two to three fold activation was observed with the TR-RIP15 chimera under a variety of conditions. However, this effect was quite modest by comparison to the more than 50 fold activation observed with TR in the presence of T3.

These transfection data suggest that both orphans may need specific ligands to activate transcription. A number of compounds considered potential ligands for orphan receptors were tested, including several hydroxycholesterols, dehydroepiandrosterone (DHEA), α-tocopherol, thyroid hormone (T3), reverse T3, and several retinoids. No specific activity was observed with any of these compounds.

In cotransfections of the orphans with RXRα, basal expression was unaffected (Figure 9). In the presence of 9-*cis*-RA, cotransfection with RXR alone resulted in strong activation of the βRARE reporter (Figure 9). Previous results (Zhang et al., Nature 358:587-591, 1992)

suggested that this effect was largely mediated by RXR homodimers, although heterodimers of RXR with endogenous RARs could also contribute. Cotransfection with RIP14-1 decreased 9-*cis*-RA induced expression by approximately 90%, and contransfection with RIP15 blocked it completely (Figure 9). RIP14-2, which bound βRARE with somewhat lower apparent affinity, decreased the level of 9-*cis*-RA induced expression by 60% (Figure 9). These inhibitory effects could be due either to direct binding of inactive RIP/RXR heterodimers to βRARE, or to indirect effects of sequestration of RXR in complexes. In either case, the results confirmed that RIP proteins were expressed in such transfections and suggested that both orphans may be involved in the complex retinoid response.

RIP110 and RIP13

RIP110 and RIP13 cDNAs were sequenced by standard techniques and deduced amino acid sequences determined, also by standard techniques. These sequences are presented in Figures 10 and 11.

The materials and methods used in the above experiments are now described.

Strains and Plasmids

LexA fusion proteins were expressed from derivatives of the LexA fusion vector (LexA(1-202)+PL) (Gyuris et al., Cell 75:791-803, 1993), which expresses the intact LexA protein. The LexA-RXR and LexA-TR fusions included human RXRα and rat TRβ sequences extending from the C-terminal portion of the DNA binding domain to the C-terminus. Analogous LexA fusions to RAR, MB67, and GR were constructed using PCR (polymerase chain reaction). For LexA-RAR, an additional fusion to the intact RARα was also generated. B42 fusion proteins were either isolated from the cDNA library as described below or inserted into a derivative of the vector pJG4-5 (Gyuris et al., Cell 75:791-803, 1993) using standard

procedures. For *in vitro* translation, appropriate fragments were cloned into a previously described bacteriophage T7 promoter expression vector (Carter et al., Mol. Cell. Biol. in press., 1994) and expressed using bacteriophage T7 RNA polymerase. Mammalian expression vectors were derivatives of CDM (Seed, Nature 329:840-842, 1987), and reporter plasmids were derivatives of pTKluc (Carter et al., Mol. Cell. Biol. in press, 1994) in which the herpes virus TK promoter directs expression of luciferase.

Yeast strains were derivatives of EGY48 (MATα leu2 trp1 ura3 his3 LEU2::pLexop6-LEU2 (ΔUAS LEU2)) (Gyuris et al., Cell 75:791-803, 1993; Zervos et al., Cell 72:223-232, 1993), in which expression of the chromosomal LEU2 gene is under the control of LexA operators. EGY48 was successively transformed with 8H18-34 (Gyuris et al., Cell 75:791-803, 1993), in which expression of the E. coli *lacZ* (β-galactosidase) gene is also under the control of *lexA* operators (selection for $URA3^+$), and derivatives of the LexA fusion expression vector LexA(1-202)+PL (Gyuris et al., Cell 75:791-803, 1993) (selection for $HIS3^+$).

cDNA Library Screening and Characterization of RXR Interactors

An oligo(dT)-primed mouse liver cDNA library was constructed using standard procedures (Ausubel et al., Current Protocols in Molecular Biology, Greene Pub. Assoc. New York, 1994) in the plasmid cgatrp2 (selection for $TRP1^+$), a derivative of the B42 expression vector pJG4-5. This plasmid also contains a tRNA suppressor gene, *supF*, which can be used to rescue only cgatrp2 (and not the other two plasmids present in the yeast host) after transformation to an *Escherichia coli* strain containing a P3 plasmid. The library was amplified and used to transform an EGY48 derivative expressing LexA-

RXR. $3x10^6$ primary yeast transformants were obtained from glucose-Ura$^-$His$^-$Trp$^-$ plates and recovered as described (Gyuris et al., Cell 75:791-803, 1993). $2x10^7$ of these cells were plated on galactose-Ura$^-$His$^-$Trp$^-$Leu$^-$ plates with or without 9-*cis*-RA. About 100 LEU2 expressing colonies were selected for X-gal testing on galactose-Ura$^-$His$^-$Trp plates. Forty were chosen for further analysis based on stable galactose dependency of both growth on Leu$^-$ plates and expression of β-galactosidase (Gyuris et al., Cell 75:791-803, 1993; Zervos et al., Cell 72:223-232, 1993). The cDNA plasmids were recovered by transformation of *E. coli* MC1063/P3 and reintroduced into host strains expressing LexA-RXR, LexA alone, or other chimeras such as LexA-Cdc2 (Gyuris et al., Cell 75:791-803, 1993) to test specificity of the interaction. Candidates that interacted specifically with LexA-RXR were selected and sequenced with a primer from the fusion site of the B42 transcription domain by the standard dideoxynucleotide method. Based on sequence information and pattern of restriction endonuclease digestion, candidate clones were divided into several classes. In some cases, further sequence information was obtained. The obtained sequences were used to search sequence databases. To isolate clones containing the full length RIP14 and RIP15 cDNAs, a mouse liver cDNA library constructed in the CDM8 plasmid by standard procedures was screened by conventional hybridization methods with fragments of RIP14 and RIP15 which were [^{32}P]-labeled by random priming.

β-Galactosidase Assay of RXR-Interacting Clones

An EGY48 derivative containing the 8H18-34 *lacZ* reporter plasmid was successively transformed with LexA and B42-fusion protein expression vectors to generate a series of strains coexpressing each LexA fusion with each B42 fusion. At least two separate colonies from glucose-

Ura$^-$His$^-$Trp$^-$ plates were selected randomly for each coexpressing strain and used to inoculate galactose-Ura$^-$ His$^-$Trp$^-$ liquid media to induce expression of the B42 fusion protein (Gyuris et al., Cell 75:791-803, 1993). Cultures were assayed for β-galactosidase as described (Ausubel et al., Current Protocols in Molecular Biology, Greene Pub. Assoc., New York, 1994).

RNA Analysis

A Northern blot containing 2μg of polyA$^+$ mRNA from the indicated tissues (Clontech, Inc., Palo Alto, CA) was hybridized with probes labeled by random priming using standard procedures (Ausubel et al., Current Protocols in Molecular Biology Greene Pub. Assoc., New York, 1994).

Cell Culture and Transfections

HepG2 cells were propagated in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum. Transfections were carried out using calcium phosphate precipitation in the same medium, or in medium supplemented with charcoal stripped serum, as described (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994). HepG2 cells plated on six well culture plates were cotransfected with 1μg of plasmids expressing intact RIPs, with or without 0.25μg of RXRα vector, and with 1.5μg of a reporter plasmid containing three copies of the βRARE (Baes et al., Mol. Cell. Biol. 14:1544-1552, 1994), and 2μg of pTKGH as an internal control. Luciferase activity was normalized using the level of growth hormone expressed from pTKGH. Each transfection was done in duplicate.

Proteins and Gel Shift Assay

RIP14 and RIP15 proteins were produced by *in vitro* translation (Promega TNT, Madison, WI) using expression vectors containing the RIP genes following a T7 promoter. To generate full length RIP14-2 constructs, fragments from RIP14-1 clone No. 15 containing regions from the

middle of the E domain to the 3′ end replaced the corresponding regions of RIP14-2 Clones No. 3 or 12, respectively. Sequences were confirmed by DNA sequencing. Human RXRα protein was expressed in *E. coli* using a bacterial expression vector based on the bacteriophage T7 promoter (Carter et al., Mol. Cell. Biol. in press., 1994). The oligonucleotides used for gel shift assays were as following: β-RARE, 5′ gatccgggtagGGTTCAccgaaAGTTCActcga 3′ (SEQ ID NO: 11); hsp27, 5′ctagacaagGGTTCAaTGCACTtgtccatcg 3′ (SEQ ID NO: 12). Hexamers that match the AGGTCA (SEQ ID NO: 13) consensus half site or its complement are capitalized. Double stranded oligonucleotides were end-labeled using [^{32}P]ATP and kinase, and free nucleotide was removed by gel filtration. Proteins were preincubated with 20μl of gel shift assay buffer (10mM Tris (pH 8.0), 40mM KCl, 0.05% NP-40, 10% glycerol, 1mM DTT, 2.5mM $MgCl_2$ and 5ng of poly dI-dC) for 10 minutes in ice. This mixture was then combined with the indicated labeled probe and incubated for 20 minutes at room temperature. Specific or nonspecific competitor oligomers were added with the probe. The mixtures were analyzed by 6% non-denaturing polyacrylamide gel electrophoresis using 0.5X Tris-Borate-EDTA (TBE) buffer at 4°C.

Identification of Ligands which Bind RXR-Interacting Proteins

Isolation of cDNAs encoding RXR-interacting proteins enables the identification and isolation of their ligands.

Accordingly, one aspect of the invention features a screening assay for the identification of compounds which specifically bind to the RXR-interacting proteins described herein. Such an assay may be carried out using a recombinant RXR-interacting protein.

In one example, the RXR-interacting protein component is produced by a cell that naturally produces substantially none of the protein or by a cell which produces functionally deficient protein; suitable cells are, e.g., those discussed above with respect to the production of recombinant receptor, most preferably, mammalian cells such as HepG2 cells. Host cells are transfected with (1) a vector which expresses a nucleic acid encoding the RXR-interacting protein (i.e., the "producer vector") and (2) a vector which includes an RXR-interacting protein binding site (e.g., for RIP14 and RIP15, the βRARE sequence described herein) positioned upstream of a target gene which may be assayed (e.g., a CAT gene, a luciferase gene, or a β-galactosidase gene) (i.e., the "reporter vector"). Using a standard transactivation assay procedure (for example, the assay described herein), RXR-interacting protein activity is assayed by measuring binding site-dependent target gene expression. Useful ligands are identified as those compounds which, when added to the host cell medium, effect a change in RXR-interacting protein-directed gene expression (as detected using any reporter vector); useful ligands according to the invention may either increase or decrease RXR-interacting protein activity.

Any suitable transactivation technique, producer vector, and binding site-containing reporter vector may be used. Descriptions of transactivation assays and generally useful vectors for the identification of ligands which bind other nuclear hormone receptors are described, e.g., in Evans et al. (U.S. Pat. No. 4,981,784, 1991); Evans et al. (WO 90/07517); Evans et al. (WO90/01428); and WO88/03168; all hereby incorporated by reference. RXR-interacting proteins which may be used to screen for ligands include wild-type molecules as well

as any appropriate chimeric protein, for example, those chimeric proteins described herein.

Candidate ligands may be purified (or substantially purified) molecules or the ligand may be one component of a mixture of ligands (e.g., an extract or supernatant obtained from cells; Ausubel et al., supra). In a mixed ligand assay, the RXR-interacting protein ligand is identified by testing progressively smaller subsets of the ligand pool (e.g., produced by standard purification techniques, e.g., HPLC or FPLC) until a single ligand is finally demonstrated to modulate the protein's activity. Candidate ligands include peptide as well as non-peptide molecules.

Alternatively, a ligand may be identified by its ability to bind an RXR-interacting protein using affinity chromatography. Recombinant protein is purified by standard techniques, from cells engineered to express the protein (e.g., those described above); the recombinant protein immobilized on a column (e.g., a Sepharose column or a streptavidin-agarose column by the immunoaffinity method of Ausubel et al., supra) and a solution containing one or more candidate ligands is passed through the column. Such a solution (i.e., such a source of candidate ligands) may be, e.g., a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured and into which the cells have secreted factors (e.g., growth factors) during culture; again, candidate ligands include peptide as well as non-peptide molecules. A ligand specific for a recombinant RXR-interacting protein is immobilized on the column (because of its interaction with the protein). To isolate the ligand, the column is first washed to remove non-specifically bound molecules, and the ligand of interest is then released from the column and collected.

Ligands isolated by the above methods (or any other appropriate method) may, if desired, be further purified (e.g., by high performance liquid chromatography; see above). Once isolated in sufficiently-purified form, a novel peptide ligand may be partially sequenced (by standard amino acid sequencing techniques). From this partial amino acid sequence, a partial nucleic acid sequence is deduced which allows the preparation of primers for PCR cloning of the ligand gene (e.g., by the method of Ausubel et al., supra).

Identification of RXR-Interacting Protein DNA Binding Sites

Identification of the RXR-interacting proteins facilitates identification of their DNA binding sites. According to one approach, DNA binding sites may be identified using a gel shift assay, e.g., as described above for the identification of the RIP14 and RIP15 binding sites. Alternatively, a transactivation assay may be utilized. Briefly, candidate DNA binding sites are inserted upstream of a target gene whose expression may be assayed and the ability of an RXR-interacting protein to bind the DNA site is assayed as its ability to activate downstream gene expression.

Alternatively, a DNA binding site may be identified by selectively retaining a protein-bound DNA fragment on a nitrocellulose filter. This approach relies on the ability of nitrocellulose to bind proteins but not double-stranded DNA. Purified RXR-interacting protein (e.g., purified by standard techniques from cells engineered to express the protein, e.g., those described above) is mixed with labelled double-stranded DNA (e.g., a random pool of DNA fragments) under conditions which allow interaction. After incubation, the mixture is suction-filtered through nitrocellulose, allowing unbound DNA to pass through the filter while retaining the

protein and any DNA specifically bound to it. Bound DNA fragments are then eluted from the filter and analyzed by gel electrophoresis or amplification and cloning. A detailed description of this technique is published in Ausubel et al. (supra).

Candidate DNA fragments for either approach may be derived, for example, from a randomly cleaved or sonicated genomic DNA library, a randomly generated set of oligonucleotides, and/or may be derived from known nuclear hormone response elements (see, e.g., Evans et al., WO90/11273).

Identification of RXR-interacting protein DNA binding sites facilitates a search for the presence of such sites upstream of known or yet unidentified genes (e.g., by an examination of sequences upstream of known genes or by standard hybridization screening of a genomic library with binding site probes). RXR-interacting protein-mediated transcriptional control of genes bearing the binding site upstream may then be investigated (e.g., by transactivation experiments as described above), potentially leading to the elucidation of novel RXR-interacting protein functions.

Chimeric Receptors

The functional domains of the RXR-interacting proteins may be swapped with the domains of other members of the nuclear hormone receptor family (see, e.g., Evans et al., WO 90/11273; Evans, *Science* 240:889, 1988) in order to produce receptors having novel properties. For example, fusion of an RXR-interacting protein DNA binding domain to the ligand-binding and gene activation domains of glucocorticoid receptor would confer hormonal regulation on genes downstream of RIP binding sites. Alternatively, fusion of an RXR-interacting protein DNA binding domain to a trans-repressing domain (see, e.g., Evans et al., WO90/14356) would result in repression of

the basal level of expression of genes bearing upstream RIP binding sites. Examples of receptor domains which may be included in a chimeric RIP receptor are described in Evans et al. (WO 90/15815) and in Evans et al. (*Science* 240:889, 1988). Construction of receptor fusion genes is carried out by standard techniques of molecular biology.

Dominant Negative Mutants

Mutants of RXR-interacting proteins may be generated which interfere with normal RIP activity. Such mutants are termed "dominant negative" and fall into at least two classes: (a) ones which bind to their DNA binding site (thereby interfering with the ability of wild-type RXR-interacting protein to bind the same site) but which do not activate ligand-dependent gene expression and (b) ones which heterodimerize with other receptors (e.g., RXR) but which do not promote the biological response associated with the wild-type heterodimer.

The first class of RIP dominant negative mutants include those receptor polypeptides which contain a wild-type DNA binding domain and a mutant gene activation domain. Such mutants are unable to transactivate a reporter gene even in the presence of ligand (e.g., as measured using a CAT reporter gene with an upstream βRARE and the standard methods described above) but retain the ability to bind a RIP DNA binding site (as evidenced, e.g., by DNA footprint analysis using a βRARE DNA sequence; Ausubel et al., supra).

The second class of RIP dominant negative mutants include those receptor polypeptides which contain a wild-type heterodimerization domain. Such a mutant interacts with its heterodimer partner and disrupts the partner's function. In one particular example, a dominant negative RIP-interacting protein may be overproduced (e.g., by

directing its expression from a very strong promoter); the abundant protein forms heterodimers with cellular RXR protein, soaking up available RXR and thereby preventing RXR homodimer formation as well as RXR heterodimer formation with other partner proteins (e.g., RAR, VDR, and T3R). Wild-type RXR-interacting protein may function as a dominant negative mutant if overproduced in this manner. However, a mutant RXR-interacting protein lacking gene activation function and/or a DNA binding domain is preferred.

Any of the above mutants may be generated by any method of random or site-directed DNA mutagenesis (see, e.g., Ausubel et al., supra).

Identification of Molecules that Modulate RXR-Interacting Protein Receptor Expression

Isolation of genes encoding RXR-interacting proteins also facilitates the identification of molecules which increase or decrease RIP expression. According to one approach, candidate molecules (e.g., peptide or non-peptide molecules found, e.g., in a cell extract, mammalian serum, or growth medium on which mammalian cells have been cultured) are added at varying concentrations to the culture medium of cells which express RIP mRNA. RIP expression is then measured by standard Northern blot analysis (Ausubel et al., supra) using RIP cDNA as a hybridization probe. The level of RIP expression in the presence of the candidate molecule is compared to the level measured for the same cells in the same culture medium but in the absence of the candidate molecule. A molecule which promotes an increase or decrease in RIP expression is considered useful in the invention.

RXR-Interacting Protein Expression

In general, RXR-interacting proteins according to the invention may be produced by transformation of a

suitable host cell with all or part of an RXR-interacting protein-encoding cDNA fragment (e.g., the cDNA described above) in a suitable expression vehicle.

Those skilled in the field of molecular biology will understand that any of a wide variety of expression systems may be used to provide the recombinant protein. The precise host cell used is not critical to the invention. The RXR-interacting protein may be produced in a prokaryotic host (e.g., *E. coli*) or in a eukaryotic host (e.g., *Saccharomyces cerevisiae* or mammalian cells, e.g., COS 1, NIH 3T3, or HeLa cells). Such cells are available from a wide range of sources (e.g., the American Type Culture Collection, Rockland, MD; also, see, e.g., Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1994). The method of transformation or transfection and the choice of expression vehicle will depend on the host system selected. Transformation and transfection methods are described, e.g., in Ausubel et al. (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1994); expression vehicles may be chosen from those provided, e.g., in *Cloning Vectors*: A Laboratory Manual (P.H. Pouwels et al., 1985, Supp. 1987).

One preferred expression system is the mouse 3T3 fibroblast host cell transfected with a pMAMneo expression vector (Clontech, Palo Alto, CA). pMAMneo provides: an RSV-LTR enhancer linked to a dexamethasone-inducible MMTV-LTR promotor, an SV40 origin of replication which allows replication in mammalian systems, a selectable neomycin gene, and SV40 splicing and polyadenylation sites. DNA encoding an RXR-interacting protein would be inserted into the pMAMneo vector in an orientation designed to allow expression. The recombinant RXR-interacting protein would be isolated as described below. Other preferable host cells which

may be used in conjunction with the pMAMneo expression vehicle include COS cells and CHO cells (ATCC Accession Nos. CRL 1650 and CCL 61, respectively).

Alternatively, an RXR-interacting protein is produced by a stably-transfected mammalian cell line. A number of vectors suitable for stable transfection of mammalian cells are available to the public, e.g., see Pouwels et al. (supra); methods for constructing such cell lines are also publicly available, e.g., in Ausubel et al. (supra). In one example, cDNA encoding the RXR-interacting protein is cloned into an expression vector which includes the dihydrofolate reductase (DHFR) gene. Integration of the plasmid and, therefore, the RXR-interacting protein-encoding gene into the host cell chromosome is selected for by inclusion of 0.01-300 μM methotrexate in the cell culture medium (as described in Ausubel et al., supra). This dominant selection can be accomplished in most cell types. Recombinant protein expression can be increased by DHFR-mediated amplification of the transfected gene. Methods for selecting cell lines bearing gene amplifications are described in Ausubel et al. (supra); such methods generally involve extended culture in medium containing gradually increasing levels of methotrexate. DHFR-containing expression vectors commonly used for this purpose include pCVSEII-DHFR and pAdD26SV(A) (described in Ausubel et al., supra). Any of the host cells described above or, preferably, a DHFR-deficient CHO cell line (e.g., CHO DHFR$^-$cells, ATCC Accession No. CRL 9096) are among the host cells preferred for DHFR selection of a stably-transfected cell line or DHFR-mediated gene amplification.

Once the recombinant RXR-interacting protein is expressed, it is isolated, e.g., using affinity chromatography. In one example, an anti-RXR-interacting

protein antibody (e.g., produced as described herein) may be attached to a column and used to isolate the RXR-interacting protein. Lysis and fractionation of RXR-interacting protein-harboring cells prior to affinity chromatography may be performed by standard methods (see, e.g., Ausubel et al., supra). Alternatively, an RXR-interacting protein fusion protein, for example, an RXR-interacting protein-maltose binding protein, an RXR-interacting protein-β-galactosidase, or an RXR-interacting protein-trpE fusion protein, may be constructed and used for RXR-interacting protein isolation (see, e.g., Ausubel et al., supra; New England Biolabs, Beverly, MA).

Once isolated, the recombinant protein can, if desired, be further purified, e.g., by high performance liquid chromatography (see, e.g., Fisher, *Laboratory Techniques In Biochemistry And Molecular Biology*, eds., Work and Burdon, Elsevier, 1980).

Polypeptides of the invention, particularly short RXR-interacting protein fragments, can also be produced by chemical synthesis (e.g., by the methods described in *Solid Phase Peptide Synthesis*, 2nd ed., 1984 The Pierce Chemical Co., Rockford, IL).

These general techniques of polypeptide expression and purification can also be used to produce and isolate useful RXR-interacting protein fragments or analogs (described herein).

Anti-RXR-Interacting Protein Antibodies

Human RXR-interacting proteins (or immunogenic fragments or analogues) may be used to raise antibodies useful in the invention; such polypeptides may be produced by recombinant or peptide synthetic techniques (see, e.g., *Solid Phase Peptide Synthesis*, supra; Ausubel et al., supra). The peptides may be coupled to a carrier protein, such as KLH as described in Ausubel et al.,